WHAT IS CLAIMED IS:

1.   A computer-implemented method for analyzing user search queries, the method comprising the acts of:

     grouping a set of previous queries into a plurality of subsets along a dimension;

     for each of the subsets of the previous queries, generating a concept network, each concept network including a plurality of units and a plurality of relationships defined between the units, wherein each unit of each concept network has a frequency weight;

     selecting one of the units; and

     constructing a histogram vector for the selected unit, the histogram vector having an element corresponding to each of the concept networks, wherein each element of the histogram vector has a value representative of the frequency weight of the selected unit in the corresponding one of the concept networks.

2.   The method of claim 1, wherein the dimension is a time dimension.

3.   The method of claim 1, wherein the dimension is defined by reference to one or more demographic characteristics of users.

4.   The method of claim 1, wherein the dimension is a geographic dimension.

5.   The method of claim 1, wherein the dimension is a vertical dimension representing a user context of the query.

6.   The method of claim 1, further comprising the act of storing the selected unit in a unit dictionary in association with the histogram vector.

7.   The method of claim 6, further comprising the acts of:

     receiving a subsequent query;

     parsing the subsequent query into one or more constituent units;

     obtaining the histogram vector for at least one of the constituent units from the unit dictionary; and

     responding to the subsequent query based at least in part on the histogram vector.

1   8.   The method of claim 7, wherein the act of responding includes suggesting
2   a related search based at least in part on the histogram vector.

1   9.   The method of claim 7, wherein the act of responding includes resolving
2   an ambiguity of one of the constituent units based at least in part on the histogram vector.

1   10.   The method of claim 6, further comprising the acts of:
2   identifying a group of units that have similar histogram vectors; and
3   storing group membership information for the units of the group in the unit
4   dictionary.

1   11.   The method of claim 6, further comprising the acts of:
2   selecting a base unit from the unit dictionary;
3   identifying a plurality of related units for the base unit in the unit dictionary;
4   determining a most common histogram vector among the related units; and
5   storing the most common histogram vector in the unit dictionary as a proxy
6   histogram vector for the base unit.

1   12.   The method of claim 11, wherein the related units include extensions of
2   the base unit.

1   13.   The method of claim 1, wherein each element of the histogram vector has
2   a binary value indicating a presence or absence of the target unit in the corresponding concept
3   network.

1   14.   The method of claim 1, further comprising the act of normalizing each of
2   the histogram vectors.

1   15.   A system for processing queries, the system comprising:
2   a concept network builder module configured to receive a set of previous user
3   queries and to generate a concept network therefrom, the concept network including a plurality
4   of units and a plurality of relationships defined between the units, wherein each unit of the
5   concept network has a frequency weight; and

31

6   a histogram builder module configured to receive a plurality of concept networks

7 generated by the concept network builder from different sets of previous user queries and further

8 configured to select one of the units and to generate a histogram vector for the selected unit,

9   wherein the histogram vector has an element corresponding to each of the concept

10 networks, wherein each element of the histogram vector has a value representative of the

11 frequency weight of the unit in the corresponding one of the concept networks.

1   16. The system of claim 15, further comprising a unit dictionary configured to

2 store the selected unit in association with the histogram vector generated for the selected unit by

3 the histogram builder module.

1   17. The system of claim 16, further comprising a histogram analysis module

2 configured to obtain units and the histogram vectors for those units from the unit dictionary and

3 to detect a pattern of the histogram vectors.

1   18. The system of claim 17, wherein the histogram analysis module is further

2 configured to select a plurality of units from the unit dictionary, to sort the selected units based

3 on the histogram vectors, and to define a group of units that have similar histogram vectors.

1   19. The system of claim 17, wherein the histogram analysis module is further

2 configured to select a base unit from the unit dictionary, to identify a plurality of related units for

3 the base unit in the unit dictionary, to determine a most common histogram vector among the

4 related units, and to store the most common histogram vector in the unit dictionary as a proxy

5 histogram vector for the base unit.

1   20. The system of claim 16, further comprising a query response module

2 configured to receive a subsequent query including one or more constituent units and to respond

3 to the subsequent query based at least in part on a histogram vector stored in the unit dictionary

4 for at least one of the constituent units.

1   21. A computer program product comprising a computer readable medium

2 encoded with program code, the program code including:

3   program code for grouping a set of previous queries into a plurality of subsets

4 along a dimension;

5   program code for generating a concept network for each of the subsets of the

6   previous queries, each concept network including a plurality of units and a plurality of

7   relationships defined between the units, wherein each unit of each concept network has a

8   frequency weight;

9   program code for selecting one of the units; and

10   program code for constructing a histogram vector for the selected unit, the

11   histogram vector having an element corresponding to each of the concept networks, wherein

12   each element of the histogram vector has a value representative of the frequency weight of the

13   selected unit in one of the concept networks.

1   22.   The computer program product of claim 21, wherein the program code

2   further includes program code for storing the selected unit in a unit dictionary in association with

3   the histogram vector.

1   23.   The computer program product of claim 21, wherein the program code

2   further includes:

3   program code for receiving a subsequent query;

4   program code for parsing the subsequent query into one or more constituent units;

5   program code for obtaining the histogram vector for at least one of the constituent

6   units from the unit dictionary; and

7   program code for responding to the subsequent query based at least in part on the

8   histogram vector.